

**CULTIVAR SPECIFICITY GENE FROM THE RICE PATHOGEN
MAGNAPORTHE GRISEA, AND METHODS OF USE**

Pursuant to 35 U.S.C. §202(c), it is acknowledged that the U.S. Government has certain rights in the invention described herein, which was made in part with funds from the United States Department of Agriculture, Grant Nos. 58-3655-3-107 and 58-3655-4-140, and from the National Institutes of Health, Grant No. GM33716-08S1.

FIELD OF THE INVENTION

This invention relates to the field of disease resistance in plants. In particular, the invention provides a novel avirulence gene from the rice blast pathogen, *Magnaporthe grisea*, and methods of using the gene and its encoded products for improving resistance of rice to this pathogen.

BACKGROUND OF THE INVENTION

Rice is a major staple food for about two-thirds of the world's population. More than ninety percent of the world's rice is grown and consumed in developing countries. Rice blast disease, caused by the fungus *Magnaporthe grisea*, threatens rice crops worldwide. The disease can cause yield losses of ten to thirty percent in infested fields. Rice blast has been an ongoing problem in rice growing areas of the southern United States. It has now become a significant problem in rice growing areas of California, as well.

It is known that there exists a relationship between the virulence of a plant pathogen and cultivars of its host

species. The gene-for-gene hypothesis has been found applicable to many host-pathogen interactions, including that of the rice blast fungus, *Magnaporthe grisea*, and its host, *Oryza sativa*. To be able to understand and
5 manipulate this host-pathogen relationship is of great practical interest as *M. grisea* is rapidly able to overcome new disease resistance in rice soon after their deployment. Moreover, *M. grisea* exists as a complex genus with many subspecific groups that are infertile, but
10 differ in their host range. How these different subspecific groups interrelate evolutionarily is of great concern to plant breeders since some of these alternate hosts are frequently found growing in close proximity to, or in rotation with rice, and *M. grisea* isolates infecting
15 these alternate hosts can sometimes also infect rice.

Gene-for-gene resistance (also known as hypersensitive resistance (HR) or race-specific resistance) depends for its activation on specific recognition of the invading pathogen by the plant. Many
20 individual plant genes have been identified that control gene-for-gene resistance. These genes are referred to as resistance (R) genes. The function of a particular R gene depends on the genotype of the pathogen. A pathogen gene is referred to as an Avr gene if its expression
25 causes the pathogen to produce a signal that triggers a strong defense response in a plant having a corresponding R gene. This response is not observed in the absence of either the Avr gene in the pathogen or the corresponding R gene in the plant. It should be noted that a single

any given Avr gene and its specific R gene are involved in a host-pathogen interaction. In this instance,

resistance generally occurs as activation of a HR response, in which the cells in the immediate vicinity of the infection undergo programmed necrosis in order to prevent the further advance of the pathogen into living plant tissue. Other features of the resistance response may also include synthesis of antimicrobial metabolites or pathogen-inhibiting enzymes, reinforcement of plant cell walls in the infected area, and induction of signal transduction pathways leading to systemic acquired resistance (SAR) in the plant.

The molecular basis of host-cultivar specificity and pathogenic variability in *M. grisea* is only beginning to be elucidated with the identification, mapping and, in some instances, cloning of specific Avr genes from pathogenic isolates of *M. grisea*. For instance, AVR2-YAMO (cultivar specificity) and PWL2 (host specificity) (Valent & Chumley, pp. 3.113 - 3.134 in Rice Blast Disease (R. Zeigler, S.A. Leong, P. Teng, Eds.), Wallingford: CAB International, 1994) both function as classic avirulence genes by preventing infection of a specific cultivar or host. AVR2-YAMO encodes a 223-amino acid protein with homology to proteases, while PWL2 encodes a 145-amino acid polypeptide which is glycine-rich. Based on the predicted amino acid sequences of the proteins, both may be secreted.

Homologs of both AVR2-YAMO and PWL2 appear to be widely distributed in rice and in other grass-infecting isolates of *M. grisea*, thereby confirming that *M. grisea* isolates which do not infect rice still may

on which the isolates do not exhibit the same level of cultivar specificity as AVR2-YAMO or PWL2.

As another example of a potentially useful Avr gene, the cultivar specificity gene AVR1-CO39, which determines avirulence on rice cultivar CO39, has been identified (Valent et al., Genetics 127: 87-101, 1991) and mapped to a position on *M. grisea* chromosome 1 (Smith & Leong, Theor. Appl. Genet. 88: 901-908, 1994). A segment of chromosome 1 that appears to contain the AVR1-CO39 gene has been isolated and cloned into a cosmid vector (Leong et al., pp. 346-352 in Rice Genetics III, Proceedings of the Third Annual Rice Genetics Symposium, G.S. Khush, Ed., Island Harbor Press, Manila, 1996); however, the gene itself heretofore has not been identified and characterized.

The availability of cloned cultivar and host specificity genes from *M. grisea* and, ultimately, the corresponding R genes from rice provides useful tools for manipulating and augmenting resistance to this pathogen in the field. Accordingly, it is an object of the present invention to provide a new cloned *M. grisea* cultivar specificity gene, AVR1-CO39, and its functional homologs for such use.

SUMMARY OF THE INVENTION

According to one aspect of the invention, there is provided an isolated nucleic acid, AVR1-CO39, from *Magnaporthe grisea* that confers rice cultivar CO39-specific avirulence to fungal plant pathogens that contain the nucleic acid. The nucleic acid preferably comprises part or all of Sequence I.D. No. 1, or

According to another aspect of the invention, there is provided a polypeptide encoded by part or all of

the isolated nucleic acid of claim 1. Preferably, the polypeptide is selected from the group of polypeptides encoded by ORFS 1, 2, 3, 4, 5, 6 and 7, corresponding to Sequence ID No's. 2, 3, 4, 5, 6, 7 and 8, respectively, and most preferably is encoded by ORF 3.

According to another aspect of the invention, a transgenic epiphytic bacterium is provided, which expresses a portion of an AVR1-CO39 gene effective to confer rice cultivar CO39-specific avirulence to microorganisms expressing the gene. Preferably, the transgenic epiphytic bacterium expresses ORF3 of Sequence ID No. 1, or a functional equivalent.

According to another aspect of the invention, a method of enhancing the scope of resistance of rice cultivar CO39 plants to pathogenic microorganisms is provided. The method comprises treating the plants with an epiphytic bacterium that expresses a portion of an AVR1-CO39 gene effective to trigger expression of a CO39-specific R gene in the plants.

According to another aspect of the invention, a second method of enhancing the scope of resistance of rice cultivar CO39 plants to pathogenic microorganisms is provided. This method comprises treating the plants with a protein extract comprising polypeptides produced by expression of AVR1-CO39, in an amount effective to trigger expression of a CO39-specific R gene in the plants.

These and other features and advantages of the present invention will be described in greater detail in

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

Various terms relating to the biological molecules of the present invention are used hereinabove and also throughout the specifications and claims. The
5 terms "substantially the same," "percent similarity" and "percent identity" are defined in detail below.

With reference to nucleic acids of the invention, the term "isolated nucleic acid" is sometimes
10 used. This term, when applied to DNA, refers to a DNA molecule that is separated from sequences with which it is immediately contiguous (in the 5' and 3' directions) in the naturally occurring genome of the organism from which it was derived. For example, the "isolated nucleic
15 acid" may comprise a DNA molecule inserted into a vector, such as a plasmid or virus vector, or integrated into the genomic DNA of a procaryote or eucaryote. An "isolated nucleic acid molecule" may also comprise a cDNA molecule.

With respect to RNA molecules of the invention
20 the term "isolated nucleic acid" primarily refers to an RNA molecule encoded by an isolated DNA molecule as defined above. Alternatively, the term may refer to an RNA molecule that has been sufficiently separated from RNA molecules with which it would be associated in its
25 natural state (i.e., in cells or tissues), such that it exists in a "substantially pure" form (the term "substantially pure" is defined below).

With respect to protein, the term "isolated protein" or "isolated and purified protein" is sometimes

used. Alternatively, this term may refer to a protein which has been sufficiently separated

from other proteins with which it would naturally be associated, so as to exist in "substantially pure" form.

The term "substantially pure" refers to a preparation comprising at least 50-60% by weight the
5 compound of interest (e.g., nucleic acid, oligonucleotide, protein, etc.). More preferably, the preparation comprises at least 75% by weight, and most preferably 90-99% by weight, the compound of interest. Purity is measured by methods appropriate for the
10 compound of interest (e.g. chromatographic methods, agarose or polyacrylamide gel electrophoresis, HPLC analysis, and the like).

With respect to antibodies of the invention, the term "immunologically specific" refers to antibodies
15 that bind to one or more epitopes of a protein of interest, but which do not substantially recognize and bind other molecules in a sample containing a mixed population of antigenic biological molecules.

With respect to oligonucleotides, the term
20 "specifically hybridizing" refers to the association between two single-stranded nucleotide molecules of sufficiently complementary sequence to permit such hybridization under pre-determined conditions generally used in the art (sometimes termed "substantially
25 complementary"). In particular, the term refers to hybridization of an oligonucleotide with a substantially complementary sequence contained within a single-stranded DNA or RNA molecule of the invention, to the substantial exclusion of hybridization of the oligonucleotide with

The term "inoculation" refers to the inoculation of a plant with a pathogen.

The term "disease defense response" refers to a change in metabolism, biosynthetic activity or gene expression that enhances the plant's ability to suppress the replication and spread of a microbial pathogen (i.e., to resist the microbial pathogen). Examples of plant disease defense responses include, but are not limited to, production of low molecular weight compounds with antimicrobial activity (referred to as phytoalexins) and induction of expression of defense (or defense-related) genes, whose products include, for example, peroxidases, cell wall proteins, proteinase inhibitors, hydrolytic enzymes, pathogenesis-related (PR) proteins and phytoalexin biosynthetic enzymes, such as phenylalanine ammonia lyase and chalcone synthase. Such defense responses appear to be induced in plants by several signal transduction pathways involving secondary defense signaling molecules produced in plants. Agents that induce disease defense responses in plants include, but are not limited to: (1) microbial pathogens, such as fungi, bacteria and viruses; (2) microbial components and other defense response elicitors, such as proteins and protein fragments, small peptides, β -glucans, elicitors and harpins, cryptogein and oligosaccharides; and (3) secondary defense signaling molecules produced by the plant, such as salicylic acid, H_2O_2 , ethylene and jasmonates.

The term "promoter region" refers to the 5' regulatory regions of a gene.

The term "reporter gene" refers to genetic

regional genes. The gene is regulated by the promoter and expression of the transgene is readily assayed.

The cloning of an AVR1-CO39 gene from *M. grisea* strain 2539 and analysis of the gene are described in detail in Example 1. The gene contains four open reading frames, one of which (ORF3) appears to play the most key role in
5 conferring cultivar specific avirulence to *Magnaporthe* isolates that carry the gene. Homologs of the strain 2539 isolate AVR1CO39 gene have been identified in a diverse array of other *Magnaporthe* isolates.

A genomic clone of AVR1-CO39 from *M. grisea*
10 strain 2539, an exemplary AVR1-CO39 of the invention, is described in detail herein and its nucleotide sequence is set forth in Example 1 as Sequence I.D. No. 1. Sequence I.D. No. 1 contains four open reading frames. It is believed that one or more of these open reading frames
15 are responsible for conferring avirulence on cultivar CO39, either by virtue of the gene product expressed from the open reading frame or by possession of critical transcription or translation regulatory sequences (see Example 1).

20 Although a genomic clone of AVR1-CO39 from *M. grisea* isolate 2539 is described and exemplified herein, this invention is intended to encompass nucleic acid sequences and proteins from other *Magnaporthe* isolates that are sufficiently similar to be used instead of the
25 isolate 2539 AVR1-CO39 nucleic acid and proteins for the purposes described below. These include, but are not limited to, allelic variants and natural mutants of Sequence I.D. No. 1, which are likely to be found in any given population of *Magnaporthe* isolates. Because such

provides an isolate AVR1-CO39 having at least about 60% (preferably 70% and more

preferably over 80%) sequence homology in the coding regions with the nucleotide sequence set forth as Sequence I.D. No. 1 (and, most preferably, specifically comprising the coding region of sequence I.D. No. 1).

5 This invention also provides isolated polypeptide products of the open reading frames of Sequence I.D. No. 1, having at least about 60% (preferably 70% or 80% or greater) sequence homology with the amino acid sequences of Sequence I.D. No's. 2, 3, 4, 5, 6 or 7, respectively.

10 Because of the natural sequence variation likely to exist among AVR1-CO39 genes, one skilled in the art would expect to find up to about 30-40% nucleotide sequence variation, while still maintaining the unique properties of the AVR1-CO39 gene and encoded polypeptides of the

15 present invention. Such an expectation is due in part to the degeneracy of the genetic code, as well as to the known evolutionary success of conservative amino acid sequence variations, which do not appreciably alter the nature of the encoded protein. Accordingly, such

20 variants are considered substantially the same as one another and are included within the scope of the present invention.

For purposes of this invention, the term "substantially the same" refers to nucleic acid or amino

25 acid sequences having sequence variation that do not materially affect the nature of the protein (i.e. its structure and/or biological activity). With particular reference to nucleic acid sequences, the term "substantially the same" is intended to refer to coding

sequences which may include conservative and non-conservative
conservative substitute amino acids in the encoded

polypeptide. With reference to amino acid sequences, the term "substantially the same" refers generally to conservative substitutions and/or variations in regions of the polypeptide that do not affect structure or function. The terms "percent identity" and "percent similarity" are also used herein in comparisons among amino acid sequences. These terms are intended to be defined as they are in the UWGCG sequence analysis program (Devereaux et al., Nucl. Acids Res. 12: 387-397, 1984), available from the University of Wisconsin.

The following description sets forth the general procedures involved in practicing the present invention. To the extent that specific materials are mentioned, it is merely for purposes of illustration and is not intended to limit the invention. Unless otherwise specified, general cloning procedures, such as those set forth in Sambrook et al., Molecular Cloning, Cold Spring Harbor Laboratory (1989) (hereinafter "Sambrook et al.") or Ausubel et al. (eds) Current Protocols in Molecular Biology, John Wiley & Sons (1998) (hereinafter "Ausubel et al.") are used.

A. Preparation of AVR1-CO39 Nucleic Acid Molecules, encoded Polypeptides and Antibodies Specific for the Polypeptides

1. Nucleic Acid Molecules

AVR1-CO39 nucleic acid molecules of the invention may be prepared by two general methods: (1) they may be synthesized from appropriate nucleotide

sequences.

The availability of nucleotide sequence

information, such as the genomic isolate having Sequence I.D. No. 1, enables preparation of an isolated nucleic acid molecule of the invention by oligonucleotide synthesis. Synthetic oligonucleotides may be prepared by
5 the phosphoramidite method employed in the Applied Biosystems 38A DNA Synthesizer or similar devices. The resultant construct may be purified according to methods known in the art, such as high performance liquid chromatography (HPLC). Long, double-stranded
10 polynucleotides, such as a DNA molecule of the present invention, must be synthesized in stages, due to the size limitations inherent in current oligonucleotide synthetic methods. Thus, for example, a 1.05 kb double-stranded molecule may be synthesized as several smaller segments
15 of appropriate complementarity. Complementary segments thus produced may be annealed such that each segment possesses appropriate cohesive termini for attachment of an adjacent segment. Adjacent segments may be ligated by annealing cohesive termini in the presence of DNA ligase
20 to construct an entire 1.05 kb double-stranded molecule. A synthetic DNA molecule so constructed may then be cloned and amplified in an appropriate vector.

AVR1-CO39 genes also may be isolated from appropriate biological sources using methods known in the
25 art. In one embodiment, a genomic clone has been isolated from a *M. grisea* strain 2539 cosmid library. In an alternative embodiment, a cDNA clone comprising one or more of the open reading frames of the genomic AVR1-CO39 locus may be isolated.

Genomic clones of the AVR1-CO39 locus may be identified by using hybridization and

washing conditions of appropriate stringency. For example, hybridizations may be performed, according to the method of Sambrook et al., using a hybridization solution comprising: 5X SSC, 5X Denhardt's reagent, 1.0% SDS, 100 µg/ml denatured, fragmented salmon sperm DNA, 0.05% sodium pyrophosphate and up to 50% formamide. Hybridization is carried out at 37-42°C for at least six hours. Following hybridization, filters are washed as follows: (1) 5 minutes at room temperature in 2X SSC and 1% SDS; (2) 15 minutes at room temperature in 2X SSC and 0.1% SDS; (3) 30 minutes-1 hour at 37°C in 2X SSC and 0.1% SDS; (4) 2 hours at 45-55°C in 2X SSC and 0.1% SDS, changing the solution every 30 minutes. Alternatively, a modification of the Amasino hybridization protocol (Anal. Biochem. 152: 304-307) is preferred for use in the present invention and is described in greater detail in Example 1.

One common formula for calculating the stringency conditions required to achieve hybridization between nucleic acid molecules of a specified sequence homology (Sambrook et al., 1989):

$$T_m = 81.5^\circ\text{C} - 16.6\log[\text{Na}^+] - 0.41(\% \text{G} + \text{C}) - 0.63(\% \text{formamide}) - 600/\text{bp in duplex}$$

As an illustration of the above formula, using $[\text{Na}^+] = [0.368]$ and 50% formamide, with GC content of 42% and an average probe size of 200 bases, the T_m is 57°C. The T_m of a DNA duplex decreases by 1 - 1.5°C with every 1% decrease in homology. Thus, targets with greater than 90% homology to the probe could be observed using a

variant of the above formula. In a preferred embodiment, clones are maintained in plasmid

cloning/expression vector, such as pGEM-T (Promega Biotech, Madison, WI) or pBluescript (Stratagene, La Jolla, CA), either of which is propagated in a suitable *E. coli* host cell.

5 AVR1-CO39 nucleic acid molecules of the invention include cDNA, genomic DNA, RNA, and fragments thereof which may be single- or double-stranded. Thus, this invention provides oligonucleotides (sense or antisense strands of DNA or RNA) having sequences capable
10 of hybridizing with at least one sequence of a nucleic acid molecule of the present invention, such as selected segments of the cDNA having Sequence I.D. No. 1. Such oligonucleotides are useful as probes for detecting AVR1-CO39 genes or mRNA in test samples of fungal isolates,
15 e.g. by PCR amplification, or for the positive or negative regulation of expression of AVR1-CO39 genes at or before translation of the mRNA into proteins.

2. Proteins

20 The AVR1-CO39 genomic isolate described herein contains four open reading frames (ORF's 1-4), whose deduced amino acid sequences are set forth herein as Sequence I.D. No's. 2-5, respectively. Any one of these polypeptides may be prepared in a variety of ways,
25 according to known methods. If produced *in situ* the polypeptides may be purified from appropriate sources, e.g., fungal isolates.

 Alternatively, the availability of nucleic acid molecules encoding the polypeptides enables production of
30 the proteins using *in vitro* expression methods known in

the art, such as cell free translation in a suitable cell free translation

system, such as wheat germ or rabbit reticulocytes. In vitro transcription and translation systems are commercially available, e.g., from Promega Biotech, Madison, Wisconsin or BRL, Rockville, Maryland.

5 According to a preferred embodiment, larger quantities of AVR1-CO39-encoded polypeptides may be produced by expression in a suitable procaryotic or eucaryotic system. For example, part or all of a DNA molecule, such as the cDNA having Sequence I.D. No. 1,
10 may be inserted into a plasmid vector adapted for expression in a bacterial cell (such as *E. coli*) or a yeast cell (such as *Saccharomyces cerevisiae*), or into a baculovirus vector for expression in an insect cell. Such vectors comprise the regulatory elements necessary
15 for expression of the DNA in the host cell, positioned in such a manner as to permit expression of the DNA in the host cell. Such regulatory elements required for expression include promoter sequences, transcription initiation sequences and, optionally, enhancer sequences.

20 The AVR1-CO39 polypeptide(s) produced by gene expression in a recombinant procaryotic or eucaryotic system may be purified according to methods known in the art. In a preferred embodiment, a commercially available expression/secretion system can be used, whereby the
25 recombinant protein is expressed and thereafter secreted from the host cell, to be easily purified from the surrounding medium. If expression/secretion vectors are not used, an alternative approach involves purifying the recombinant protein by affinity separation, such as by
30 immunological interaction with antibodies that bind

invention, prepared by the aforementioned methods, may be used in a variety of ways. For example, the methods for

analyzing the functional activity, i.e. ability to confer avirulence, are described in Example 1.

The present invention also provides antibodies capable of immunospecifically binding to polypeptides of the invention. Polyclonal or monoclonal antibodies directed toward any of the peptides encoded by the ORFs of AVR1-CO39 may be prepared according to standard methods. Monoclonal antibodies may be prepared according to general methods of Köhler and Milstein, following standard protocols. In a preferred embodiment, antibodies are prepared, which react immunospecifically with various epitopes of the AVR1-CO39-encoded polypeptides.

Polyclonal or monoclonal antibodies that immunospecifically interact with one or more of the polypeptides encoded by AVR1-CO39 can be utilized for identifying and purifying such proteins. For example, antibodies may be utilized for affinity separation of proteins with which they immunospecifically interact. Antibodies may also be used to immunoprecipitate proteins from a sample containing a mixture of proteins and other biological molecules. Other uses of the antibodies are described below.

B. Uses of AVR1-CO39 Nucleic Acids, Encoded Proteins and Antibodies

The potential of recombinant genetic engineering methods to enhance disease resistance in agronomically important plants has received considerable attention in recent years. Protocols are currently available for the stable introduction of genes into

... of a plant cell, or into an epiphytic microorganism, can enhance the

plant's ability to resist pathogen attack. AVR1-CO39-
encoded proteins of the invention may also be applied
directly to a plant, to induce a disease defense
response.

5

1. AVR1-CO39 Nucleic Acids

AVR1-CO39 nucleic acids (genomic clones or
cDNAs) may be used for a variety of purposes in
accordance with the present invention. The DNA, RNA, or
10 fragments thereof may be used as probes to detect the
presence of and/or expression of AVR1-CO39 genes.
Methods in which AVR1-CO39 nucleic acids may be utilized
as probes for such assays include, but are not limited
to: (1) *in situ* hybridization; (2) Southern hybridization
15 (3) northern hybridization; and (4) assorted
amplification reactions such as polymerase chain
reactions (PCR). The AVR1-CO39 nucleic acids of the
invention may also be utilized as probes to identify
homologs from other *Magnaporthe* isolates. As described
20 above, AVR1-CO39 nucleic acids are also used to advantage
to produce large quantities of substantially pure AVR1-
CO39 proteins, or selected portions thereof.

Of perhaps greater significance, however, is
the use of AVR1-CO39 nucleic acids to broaden the scope
25 of resistance of rice cultivars carrying the CO39
resistance gene to pathogens other than *M. grisea*
isolates carrying the AVR1-CO39 avirulence gene. For
instance, in one embodiment of the invention, the AVR1-
CO39 coding region is operably linked to a heterologous
30 promoter, preferably one that is either generally

promoter, or a promoter of a gene encoding a

as promoters of genes encoding lipoxygenases

(preferably from plants, most preferably from rice, e.g., Peng et al., J. Biol. Chem. 269: 3755-3761, 1994; Peng et al., Abstract presented at the general meeting of the International Program on Rice Biotechnology, Malacca, Malaysia, Sept. 15-19, 1997);

b) promoters of genes encoding peroxidases (preferably from plants, most preferably from rice, e.g., Chittoor et al., Mol. Plant-Microbe Interactions 10: 861-871, 1997);

c) promoters of genes encoding hydroxymethylglutaryl-CoA reductase (preferably from plants, most preferably from rice, e.g., Nelson et al., Plant Mol. Biol. 25: 401-412, 1994);

d) promoters of genes encoding phenylalanine ammonia lyase (preferably from rice; e.g., Lamb et al., Abstract of the general meeting of the International Program on Rice Biotechnology, Malacca, Malaysia, Sept. 15-19, 1997)

e) promoters of genes encoding glutathione-S-transferase (preferably from plants, most preferably from rice, or alternatively, the PRP1 promoter from potato);

f) promoters from pollen-specific genes, such as corn Zmg13, which has been shown to be expressed in rice transgenic pollen carrying the corn gene (Aldemita et al., Abstract of the general meeting of the International Program on Rice Biotechnology, Malacca, Malaysia, Sept. 15-19, 1997);

g) promoters from genes encoding chitinases (preferably from plants, most preferably from rice; e.g., Zhu & Lamb, Mol. Gen. Genet. 226: 239-246, 1991);

i) promoters from plant (preferably rice) viral genes, either contained on a bacterial plasmid or on a plant viral vector, as described by Hammond-Kosack et al., Mol. Plant-Microbe Interactions 8: 181-185 (1994);

5 j) promoters from genes involved in the plant (preferably rice) respiratory burst (e.g., Groom et al., Plant J. 10(3): 515-522, 1996); and

k) promoters from plant (preferably rice) anthocyanin pathway genes (e.g., Reddy, pp 341-352 in 10 Rice Genetics III, supra; Reddy et al., Abstract of the general meeting of the International Program on Rice Biotechnology, Malacca, Malaysia, Sept. 15-19, 1997).

The chimeric gene is then used to transform rice cultivars that already carry the appropriate R gene. 15 Upon wounding or challenge with a plant pathogen, the resulting transgenic plants would be induced to produce the AVR1-CO39 gene product, thereby triggering the R gene defense response. In this embodiment, care must be taken to avoid using a promoter that is induced by necrosis, 20 since use of such a promoter could result in a self-perpetuating hypersensitive response that may be lethal to the plant (see, e.g., Kim et al., Proc. Natl. Acad. Sci. USA 91: 10445-10449, 1994).

In a preferred embodiment, a coding region of 25 AVR1-CO39 (preferably the coding region corresponding to ORF3) is inserted into an expression vector in a microorganism that grows epiphytically on rice plants. A suspension of such recombinant microorganisms is sprayed on rice cultivars carrying the appropriate R gene. Upon 30 pathogen attack, two levels of protection can occur: (1)

1. a high level of protection is achieved by the gene product produced by

the recombinant epiphytes is carried into the plant tissue at the wound site, where it interacts with the corresponding R gene product and induces an internal disease defense response. Thus, this pre-treatment
5 confers resistance to *Maganporthe* isolates (and, presumably, other plant pathogens) which normally are virulent on those cultivars. This embodiment is described in greater detail in Example 3.

In connection with the use of epiphytic
10 bacteria, it should be noted that bacterial and phage expression and delivery systems, such as those commercially available from InVitrogen, will be particularly useful. The bacterial system expresses a protein hybrid with pilin, such that the foreign protein
15 is exposed on the exterior of the bacterium. The phage system also expresses a hybrid protein with coat component and exterior exposure of the foreign protein.

The AVR1-CO39 gene also may be used as a tool to identify and isolate its corresponding R gene. Thus,
20 in a manner similar to that described for isolation of the tomato CF-9 gene for resistance to *Cladosporium fulvum* (Jones et al., Science 266: 789-793, 1994), the R gene in rice that corresponds to AVR1-CO39 can be isolated by transposon tagging: (1) AVR1-CO39 is
25 transformed into, and constitutively expressed in a susceptible rice line; (2) the transgenic line is crossed with a resistant line that carries an identifiable transposon; (2) seedlings of F1 progeny constitutively expressing both the Avr gene and the corresponding R gene
30 should die, thereby enabling a simple screening for live

transposon. The transposon, along with the gene it has

2. AVR1-CO39 Proteins and Antibodies

Purified gene products of AVR1-CO39, or fragments thereof, may be used to produce polyclonal or monoclonal antibodies, which also may serve as sensitive
5 detection reagents for the presence and accumulation of AVR1-CO39 polypeptides in transformed microbial epiphytes, transgenic plants, or other biological materials. Polyclonal or monoclonal antibodies immunologically specific for AVR1-CO39 polypeptides may
10 be used in a variety of assays designed to detect and quantitate the proteins. Such assays include, but are not limited to: (1) flow cytometric analysis; (2) immunochemical localization of expressed proteins in cells or tissues; and (3) immunoblot analysis (e.g., dot
15 blot, Western blot) of extracts from various cells and tissues. Additionally, as described above, antibodies can be used for purification of AVR1-CO39 polypeptides (e.g., affinity column purification, immunoprecipitation).

20 In a preferred embodiment, purified AVR1-CO39 polypeptides (most preferably from ORF3) are used as a pre-treatment or co-treatment to confer broad-spectrum pathogen resistance to rice cultivars carrying the CO39 R gene. Thus, in a manner similar to the above-described
25 use of AVR1-CO39-expressing epiphytic microorganisms, a solution of the peptide is applied to the plants, and subsequent or concurrent wounding or inoculation with a pathogenic microorganism brings the peptide into contact with the R gene product, thereby stimulating a defense
30 response. The inventors have experimentally demonstrated

The following specific examples are provided to illustrate embodiments of the invention. They are not intended to limit the scope of the invention in any way.

5

EXAMPLE 1

Cloning and Analysis of AVR1-CO39

A chromosomal segment putatively containing the cultivar specificity gene, AVR1-CO39, was isolated from
10 *M. grisea* strain 2539 using a map-based cloning approach, followed by chromosome walking (Leong et al., 1996). In this example we describe the identification, cloning and analysis of the AVR1-CO39 gene.

15 Methods:

Hybridization protocol. Hybridization methods were modified from Amasino (1986) "Acceleration of Nucleic Acid Hybridization Rate by Polyethylene Glycol", Analytical Biochemistry 152:304-307. The hybridization
20 buffer was prepared according the Amasino protocol, but without the PEG and NaCl and with reduced concentrations of NaHPO₄: 0.125M NaHPO₄, 7% SDS, 50% formamide, 1.0 mM EDTA, pH 7.2. High stringency conditions were used (42 °C, 16 h). Post hybridization washes were: one rinse
25 with 2XSSC at room temperature; one wash in 2XSSC for 10 min at 65 °C; one wash in 2XSSC, 15 min at 65 °C; one final wash in 0.1 X SSC, 0.1% SDS for 15 min at 65 °C. The final washing conditions were of greater stringency than were the hybridization conditions, giving a T_m of 68.
30 Thus, greater than 95% homology would be required to maintain a hybrid. None of the post hybridization

Chromosome walking strategy.

35 DNA library of *M. grisea* strain 2539 consisting of 1,194

clones was constructed in cosmid vector pMLF1 (Leong et al., 1994, *supra*) and pMLF2 (An et al., Gene 176: 93-96, 1996). Clones were templated individually as colony blots as well as in pools in which the DNA was
5 restriction digested, electrophoresed and blotted. The latter blots were used initially to identify candidate pools containing hybridizing clones. Colony blots derived from these pools were then screened. Steps were performed using endclones prepared from the insert DNA by
10 digesting the cosmid clones with *Apa*I, which does not digest the vector, and recircularizing the plasmid by ligation. This procedure results in a derivative containing DNA from each end of the insert (An et al., 1996 *supra*). Liberation of both ends of the insert from
15 the vector was achieved by digesting with *Apa*I and *Not*I. The required endclone was then identified by virtue of its inability to hybridize with the previous cosmid in the walk.

Transformation of virulent strain Guy11 with
20 **cosmids within the AVR1-CO39 locus:** Cosmids from within the genetic interval containing AVR1-CO39 were introduced into Guy11 using the transformation protocol described in LEUNG et al. (1990). The procedure was modified as follows: After the protoplasts were incubated in
25 complete medium (CM)+sorbitol, they were poured into 100 ml molten (45°C) CM+20% sucrose agar. The agar was then poured into four petri plates. When the agar had solidified (1 h) it was overlaid with 15 ml of 1.5% water agar containing 800 µg/ml hygromycin B (300 µg/ml final
30 concentration).

...
... first two clones. The mutant 1.15 kb fragments were then

cloned into pCB1004 a hyg^r vector from J. Sweigard (Dupont). Plasmids were linearized by NotI digestion and transformed into Guy11 protoplasts.

Initial frameshift mutations were created in
5 ORFs 2 and 3 by digestion and religation as follows:

Frameshift in ORF2: The AccI site at nucleotide (nt) position 499 was cut and the 2 nt 3' overhangs were trimmed off with T4 DNA polymerase. The site was then religated resulting in the removal of 4 bp
10 or a net frameshift of -2. The nucleotide sequence changed from 5' CTAGACAGTCTACCTCTCTGCCA 3' (SEQ ID NO:9) to 5' CTAGACAGTACCTCTCTGCCA 3' (SEQ ID NO:10).

Frameshift in ORFS 2 & 3: The PflMI site at nt position 641 was cut and the 3 nt 3' overhang was trimmed
15 off with T4 polymerase. A klenow-filled HindIII fragment containing the streptomycin resistance gene cassette from pHP450 (Prentki and Kritsch, Gene 29: 303, 1984) was ligated to the flush-ended PflMI fragment. The conserved HindIII site was then digested and religated. The net
20 effect was the substitution of the 3 nt in the PflMI site with 4 nt from the HindIII site. This created a net frameshift of +1. The nucleotide sequence change was from 5' CCAGCAGCCCAATGCTTGGAAAGATTG 3' (SEQ ID NO:11) to 5' CCAGCAGCCCAAAGCTTTGGAAAGATTG 3' (SEQ ID NO:12).

25 In the AAccI construct, the peptide retains only 19 aa of its original sequence and is truncated after 36 aa. The native ORF2 peptide is 77 aa. In the APfl construct, the ORF2 peptide sequence is almost unchanged except for the terminal 10 aa and the resulting
30 peptide is 17 aa longer. ORF3, on the other hand,

Change" site-directed mutagenesis (Stratagene) using

primers designed to introduce an extra G nucleotide after the ATG:

P1: CAACGTACTAGAAATGGAGTAATAAGTACC (SEQ ID NO:13)

P2: GGTACTTATTAGTCCATTTCTAGTACGTTG (SEQ ID NO:14)

5 The mutagenesis basically abolished the ORF completely.

Creation of ATG mutations in open reading frames at the AVR1-CO39 gene locus: The 1.05 kb fragment containing AVR1-CO39 was cloned into pCB1004 and
10 Quickchange mutagenesis (Stratagene) was used to make the following mutant constructs:

ΔORF1 (ATG->TTT): Start codon of ORF1 was eliminated by quick change mutagenesis using a primer with the mutant ATG sequence.

15 **ΔORF3 (ATG->TTT):** Start codon of ORF3 was eliminated by quick change mutagenesis using a primer with the mutant ATG sequence.

 The following clone was made but the mutant allele has not yet been tested by transformation to
20 determine the phenotype.

ΔORF2 (ATG->TTT): Start codon of ORF2 was eliminated by quick change mutagenesis using a primer with the mutant ATG sequence.

25 **Results:**

 As mentioned above, a gene conferring cultivar-specific interactions with rice cultivar CO39 was isolated from *M. grisea* strain 2539 using a map-based cloning approach, followed by a 20-step chromosome walk.

 The 1.05 kb region of Chromosome 1 is set forth below

as SEQ ID NO:1:

SEQ ID NO: 1 (5' → 3')

5 GATCTGTAAA TTACATATAT TTATTTTGCC GCATTTTGCT AACCGCCTAT
TCTTTTAAA ATTTTAAAGA TTAAGAAGC AATTCAATTT TGGGTTCTAC
ACAAATTAAC AATTCGTCCA AAAGAGGTAT TTAAGCGAAG ATTTGGCATT
10 TTTTAAATCC ATTTTAAAAA AAATACATCT GCTTTAAGCC ACCTTTGCCA
AGGGTACCCG GCTAGCATAG CCTTGGTTAC CAAAAACGGC TAAAGCTGTC
GATCTATACT ACATTGGGCG CTCTGAACAA CTAAGCAACA GCGAGGAGAT
15 T5
CACACCCTAA ATCATGCTGC TAGTAATGCG ATATAATGGC CAAACAACGT
ORF1+
ACTAGAAATG ACTAATAAGT ACCCAGTCAA GTCAACTTGC TGTAGTATTA
-ORF5 ORF2+
20 TATTTAAGCA AGCGTCATT TACTGCCAGG GCAAGTTTAT CAATGGGACC
T1
AGTGTTCTCC CTCCTCTGGA CAACTCAGTT CTTTGCAAAC GCTAGACAGT
CTACCTCTCT GCCACCATTT TTACTTTTCA AAAATTTACT CCTTGCCGCT
25 T4 ORF3+
ACTGAAACTT CTACAATTGA AAGAGCCAC AATGAAAGTC CAAGCTACAT
TGGCCACCCT TATCGCCCTT GCGGCTTACT TTCCAGCAGC CAATGCTTGG
T2
30 AAAGATTGCA TCATCCAACG TTATAAAGAC GCGGATGTCA ACAACATATA
TACTGCCAAT AGGAACGAAG AGATAACTAT TGAGGAATAT AAAGTCTTCG
ORF5+ ORF4
TTAATGAGGC CTGCCATCCC TACCCAGTTA TACTTCCCGA CAGATCGGTC
35 T3
CTTCTGGCG ATTTTACATC AGCTTAGGCT GACGACGATG AGTCTTGTTC
T5 ORF7+
ATCAATAAGA GTCCAGGTTG AAAAATTGCG CACCATGGTA ATAGAGGGTT
40 ATTTATCTCG GAATAGCAGC CGTGTGTGCA ATTATCAGG CTGTTCTCT

DNA sequence analysis revealed four small open reading frames of 45, 77, 89 and 69 amino acids in length (ORF1, ORF2, ORF3, ORF4, respectively, as shown on SEQ ID No. 1 above). The amino acid sequences encoded by the
5 four open reading frames are set forth below as Sequence ID No's. 2, 3, 4 and 5, respectively. Three other open reading frames were also identified (ORFS 5, 6, and 7, set forth below as SEQ ID NOS: 6, 7 and 8, respectively.

10 **AVR1-CO39 ORF1 (SEQ ID NO: 2)**

MTNKYPVKST CCSIIFNEAS IYCQKGFING TSVLPPLDNS
VLCKR

AVR1-CO39 ORF2 (SEQ ID NO: 3)

15 MGPVFSLLWT TQFFANARQS TSLPPFLLFK NLLLAATETS
TIERAHNESP SYIRHPYRPC GLLSSSQCLE RLHHPTL

AVR1-CO39 ORF3 (SEQ ID NO: 4)

MKVQATFATL IALAAYFPAA NAWKDCIIQR YKGDVNNIY
20 TANRNEEITI EEEKVFEVNEA CHYPVILPD RSVLSGDFTS
AYADDDDESC

AVR1-CO39 ORF4 (SEQ ID NO: 5)

MAGLINEDFI FLNSYLFVPI GSIYVVDIAV FITLDDAIFE
25 SIGQWKVSRK GDKGGECSLD FHCGLFQL

AVR1-CO39 ORF5 (SEQ ID NO: 6)

MDASLNIILQ QVELTGYLLV ISSTLFGHYI ALLAA

AVR1-CO39 ORF7 (SEQ ID NO: 8)

MVIEGYLSRN SSRVCNYHGC SSAIGILEAG QIYGNSNQLS
LSMDSPVEWR RRRI (continues beyond cloned DNA)

The sequence surrounding the ATG of ORF3
5 matched four out of five of the conserved bases found in
fungal translation start sites and contained a
hydrophobic amino terminus punctuated by a lysine in
position 2, and two putative cleavage sites for removal
of the signal peptide. A fourth open reading frame
10 (ORF4) was identified on the opposite strand. However,
the sequence surrounding that ATG contained only two
matches with the fungal translation start site consensus
sequence.

Site-directed mutations in ORF1, ORF2 and ORF3
15 were created in order to assess the roles of these ORFs
in conferring avirulence. The translation start codon of
each ORF was converted from ATG to TTT. In ORFs 1 and 3,
these mutations led to loss of avirulence. Frameshift
mutations in ORF1 and ORF3 also led to a loss of
20 avirulence, while the frameshift mutation of ORF 2 did
not. Taken together, these data indicate a role for ORF1
and ORF3 in conferring avirulence to *M. grisea* strain
2539 on rice cultivar CO39.

The absence of a splice site and a lariat
25 sequence, as well as any putative TATA element
immediately upstream of the ATG of ORF1 may indicate that
ORF1 overlaps sequences critical to the promotion of
transcription of AVR1-CO39.

**Distribution of AVR1-CO39 Homologs
in Diverse Isolates of *Magnaporthe grisea***

The distribution of AVR1-CO39 homologs was
5 investigated by probing a large sample of host-specific
forms of *M. grisea* with a segment of AVR1-CO39 DNA, using
hybridization conditions such as those described in
Example 1. The results of this survey indicate that
isolates infecting rice, *Digitaria* and wheat largely lack
10 homologs of Avr-CO39. However, homologs of the gene were
commonly found in Elutine (*Setaria*)-infecting isolates.
Moreover, a detailed analysis of the AVR1-CO39 locus from
virulent rice isolate Guy 11 indicated that at least 20
Kb of DNA corresponding to and containing the AVR1-CO39
15 locus of isolate 2539 was absent.

EXAMPLE 3

**Improved Resistance to *M. grisea* Infection
in Rice Plants Sprayed with Bacterial
20 Epiphytes expressing ORF 3 of AVR1-CO39**

The ORF3 of AVR1-CO39 described in Example 1
was transferred into a pET expression vector in
25 *Escherichia coli*. A suspension containing the
transformed *E. coli* was sprayed onto leaves of rice
plants carrying the corresponding R gene for AVR1-CO39.
The plants were then inoculated with *M. grisea* isolate
Guy 11, which is a virulent strain on the plant cultivars
30 tested. As a control, plants were sprayed with an *E.*
coli suspension which did not contain the ORF-3 encoding

... number as compared to inoculated control plants pre

treated with *E. coli* lacking the ORF3-expressing plasmid. These data support the role of ORF3 in conferring avirulence in *M. grisea*.

5

EXAMPLE 4
Improved Resistance to *M. grisea* Infection
in Rice Plants Sprayed with Protein Encoded by
ORF 3 of AVR1-CO39

10 The ORF3 of AVR1-CO39 described in Example 1 was transferred into a pET expression vector in *Escherichia coli*. Protein extracts from IPTG-induced *E. coli* cells carrying either the pET vector alone (control) or the pET-ORF3 construct were tested for their effects
15 on virulence. The cellular protein extracts were concentrated by ammonium sulfate precipitation. Cultivar CO39 was inoculated with virulent *M. grisea* strain Guy11 in combination with the concentrated protein extract to give 5×10^5 conidia and 20 mg total protein extract in 10
20 ml sterile water.

 Inoculated plants co-treated with the ORF3-containing protein extract displayed reduced lesion size and number as compared to inoculated control plants co-treated with protein extract lacking ORF3. These data
25 further support the role of ORF3 in conferring avirulence in *M. grisea*.

 While certain of the preferred embodiments of the present invention have been described and
30 specifically exemplified above, it is not intended that

 be limited to the following claims.

K400v-2
The Gng Gng Gng The Gng Gng Gng The Gng Gng Gng The Gng Gng Gng The Gng Gng Gng

<211> 77
 <212> PRT
 <213> Magnaporthe grisea

<400> 3

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Gly | Pro | Val | Phe | Ser | Leu | Leu | Trp | Thr | Thr | Gln | Phe | Phe | Ala | Asn |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Ala | Arg | Gln | Ser | Thr | Ser | Leu | Pro | Pro | Phe | Leu | Leu | Phe | Lys | Asn | Leu |
| | | 20 | | | | | 25 | | | | | | 30 | | |
| Leu | Leu | Ala | Ala | Thr | Glu | Thr | Ser | Thr | Ile | Glu | Arg | Ala | His | Asn | Glu |
| | | 35 | | | | 40 | | | | | | 45 | | | |
| Ser | Pro | Ser | Tyr | Ile | Arg | His | Pro | Tyr | Arg | Pro | Cys | Gly | Leu | Leu | Ser |
| | 50 | | | | 55 | | | | | | 60 | | | | |
| Ser | Ser | Gln | Cys | Leu | Glu | Arg | Leu | His | His | Pro | Thr | Leu | | | |
| 65 | | | | | 70 | | | | | 75 | | | | | |

<210> 4
 <211> 89
 <212> PRT
 <213> Magnaporthe grisea

<400> 4

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Lys | Val | Gln | Ala | Thr | Phe | Ala | Thr | Leu | Ile | Ala | Leu | Ala | Ala | Tyr |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Phe | Pro | Ala | Ala | Asn | Ala | Trp | Lys | Asp | Cys | Ile | Ile | Gln | Arg | Tyr | Lys |
| | | 20 | | | | | 25 | | | | | | 30 | | |
| Asp | Gly | Asp | Val | Asn | Asn | Ile | Tyr | Thr | Ala | Asn | Arg | Asn | Glu | Glu | Ile |
| | | 35 | | | | 40 | | | | | | 45 | | | |
| Thr | Ile | Glu | Glu | Tyr | Lys | Val | Phe | Val | Asn | Glu | Ala | Cys | His | Pro | Tyr |
| | 50 | | | | 55 | | | | | 60 | | | | | |
| Pro | Val | Ile | Leu | Pro | Asp | Arg | Ser | Val | Leu | Ser | Gly | Asp | Phe | Thr | Ser |
| 65 | | | | | 70 | | | | | 75 | | | | | 80 |
| Ala | Tyr | Ala | Asp | Asp | Asp | Glu | Ser | Cys | | | | | | | |
| | | | | | 85 | | | | | | | | | | |

<210> 5
 <211> 68
 <212> PRT
 <213> Magnaporthe grisea

<400> 5

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Gly | Leu | Ile | Asn | Glu | Asp | Phe | Ile | Phe | Leu | Asn | Ser | Tyr | Leu |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Phe | Val | Pro | Ile | Gly | Ser | Ile | Tyr | Val | Val | Asp | Ile | Ala | Val | Phe | Ile |
| | | 20 | | | | | 25 | | | | | | 30 | | |
| Thr | Leu | Asp | Asp | Ala | Ile | Phe | Pro | Ser | Ile | Gly | Cys | Trp | Lys | Val | Ser |
| | | 35 | | | | 40 | | | | | 45 | | | | |
| Arg | Lys | Gly | Asp | Lys | Gly | Gly | Glu | Cys | Ser | Leu | Asp | Phe | His | Cys | Gly |
| | 50 | | | | 55 | | | | | | 60 | | | | |
| Leu | Phe | Gln | Leu | | | | | | | | | | | | |
| 65 | | | | | | | | | | | | | | | |

<210> 6
 <211> 6
 <212> 6
 <213> 6

Met Asp Ala Ser Leu Asn Ile Ile Leu Gln Gln Val Asp Leu Thr Gly
 1 5 10 15
 Tyr Leu Leu Val Ile Ser Ser Thr Leu Phe Gly His Tyr Ile Ala Leu
 20 25 30
 Leu Ala Ala
 35

<210> 7
 <211> 34
 <212> PRT
 <213> Magnaporthe grisea

<400> 7
 Met Arg Pro Ala Ile Pro Thr Gln Leu Tyr Phe Pro Thr Asp Arg Ser
 1 5 10 15
 Phe Leu Ala Ile Leu His Gln Leu Thr Leu Thr Thr Met Ser Leu Val
 20 25 30
 Asp Gln

<210> 8
 <211> 54
 <212> PRT
 <213> Magnaporthe grisea

<400> 8
 Met Val Ile Glu Gly Tyr Leu Ser Arg Asn Ser Ser Arg Val Cys Asn
 1 5 10 15
 Tyr His Gly Cys Ser Ser Ala Ile Gly Ile Leu Glu Ala Gly Gln Ile
 20 25 30
 Tyr Gly Asn Ser Asn Gln Leu Ser Leu Ser Met Asp Ser Pro Val Glu
 35 40 45
 Trp Arg Arg Arg Arg Ile
 50

<210> 9
 <211> 23
 <212> DNA
 <213> Magnaporthe grisea

<400> 9
 ctagacagtc taactctctg cca

23

<210> 10
 <211> 21
 <212> DNA
 <213> Magnaporthe grisea

<400> 10
 ctagacagta cctctctgac a

21

<400> 11
 ccacacagca atgcttggaa aaattg

20

<210> 12
<211> 27
<212> DNA
<213> Magnaporthe grisea

<400> 12
ccagcagcca aagctttgga aagattg 27

<210> 13
<211> 30
<212> DNA
<213> Magnaporthe grisea

<400> 13
caacgtacta gaaatggagt aataagtacc 30

<210> 14
<211> 30
<212> DNA
<213> Magnaporthe grisea

<400> 14
ggtacttatt agtccatttc tagtacgttg 30